

UNIVERSIDAD POLITÉCNICA SALESIANA
SEDE QUITO

CARRERA:
INGENIERÍA ELECTRÓNICA

Trabajo de titulación previo a la obtención del título de:
INGENIEROS ELECTRÓNICOS

TEMA:
DESARROLLO DE UN PROTOTIPO DE GAFAS PARA LECTURA DE
TEXTO CON VISIÓN ARTIFICIAL QUE ASISTA A PERSONAS CON
DISCAPACIDAD VISUAL

AUTORES:
DAVID ALEJANDRO BASANTES VARELA
ERNESTO ALEXANDER CHALACO CHAMBA

TUTOR:
GUSTAVO JAVIER CAIZA GUANOCHANGA

Quito, septiembre del 2019

CESIÓN DE DERECHOS DE AUTOR

Nosotros David Alejandro Basantes Varela, con documento de identificación N° 1721256715 y Ernesto Alexander Chalaco Chamba con documento de identificación N° 1900551639, manifestamos nuestra voluntad y cedemos a la Universidad Politécnica Salesiana la titularidad sobre los derechos patrimoniales en virtud de que somos autores del trabajo de titulación intitulado: “DESARROLLO DE UN PROTOTIPO DE GAFAS PARA LECTURA DE TEXTO CON VISIÓN ARTIFICIAL QUE ASISTA A PERSONAS CON DISCAPACIDAD VISUAL.”, mismo que ha sido desarrollado para optar por el título de: Ingenieros Electrónicos, en la Universidad Politécnica Salesiana, quedando la Universidad facultada para ejercer plenamente los derechos cedidos anteriormente.

En aplicación a lo determinado en la Ley de Propiedad Intelectual, en nuestra condición de autores nos reservamos los derechos morales de la obra antes citada. En concordancia, suscribimos este documento en el momento que hacemos entrega del trabajo final en formato impreso y digital a la Biblioteca de la Universidad Politécnica Salesiana.



.....
David Alejandro Basantes Varela

C.I. 1721256715



.....
Ernesto Alexander Chalaco Chamba

C.I. 1900551639

Quito, septiembre del 2019

DECLARATORIA DE COAUTORÍA DEL DOCENTE TUTOR

Yo declaro que bajo mi dirección y asesoría fue desarrollado el Proyecto Técnico. “DESARROLLO DE UN PROTOTIPO DE GAFAS PARA LECTURA DE TEXTO CON VISIÓN ARTIFICIAL QUE ASISTA A PERSONAS CON DISCAPACIDAD VISUAL”. Realizado por David Alejandro Basantes Varela y Ernesto Alexander Chalaco Chamba.

Obteniendo un producto que cumple con todos los requisitos estipulados por la Universidad Politécnica Salesiana, para ser considerados como trabajo final de titulación.

Quito, septiembre del 2019



.....
Gustavo Javier Caiza Guanochanga
C.I. 1721192191

DEDICATORIAS

Quiero dedicar este esfuerzo a mi familia a los que siempre han estado junto a mí y a los que se fueron antes de culminar una etapa más en mi vida. A mi tía Etelvina, que hoy no puede acompañarme, pero siempre me apoyó en todo momento y de quien aprendí que no debemos esperar de nadie nada, pero darlo todo siempre. A mi tío José, que me enseñó que el esfuerzo y la dedicación no tiene edad, mientras nos quede fuerza nuestros sueños y objetivos se pueden cumplir. A mis padres, Verónica y Agustín, cada uno a su manera, supieron inculcar en mí, con su ejemplo, el valor de la familia y la unidad, la moral y la conciencia necesaria para tomar las mejores decisiones, y también para saber sobreponerme de las caídas y obstáculos que nos da la vida. A mis hermanos, Carlos y Gabriela, quienes han estado en todo este camino, a ellos les quiero agradecer por permitirme ser parte de sus vidas.

David

Este proyecto se lo dedico con mucho cariño a mi familia; de manera especial a mis padres. A mi madre Lilia por su infinito amor y comprensión en los momentos más difíciles, a mi padre Ernesto por ser mi mejor consejero y ejemplo a seguir. A mis hermanos Paola y Lenin por compartir momentos significativos conmigo, que hoy me permiten tener recuerdos inolvidables. A mi novia Paola por siempre estar dispuesta a escucharme y acompañarme en momentos arduos de mi vida. Este trabajo y todo lo que logre hacer será gracias a ustedes.

Ernesto

AGRADECIMIENTOS

Quiero empezar agradeciendo al rector de la UPS por haberme permitido culminar mis estudios, gracias a sus acciones hoy estoy aquí con las mismas ganas y el mismo entusiasmo con el que decidí acabar mi carrera y cumplir un objetivo más en mi vida. A la Universidad por mantener siempre un estándar de calidad adecuado en cuanto a su infraestructura; laboratorios, aulas, equipos de primera, que me permitió desarrollar mis habilidades y potenciar de mejor manera los proyectos que semestre a semestre se presentaban.

A las y los ingenieros de la institución que, con su esfuerzo y dedicación, transmitieron no solo su conocimiento científico, sino también sus experiencias personales y laborales, que me permitieron entender cómo funciona el mundo y a que nos enfrentamos, gracias por permitirme aprender hasta el último día que pude compartir un aula de clases con ustedes.

Por último, quiero dar gracias a mis padres quienes han compartido y han luchado hombro a hombro junto a mí para salir adelante en esta etapa tan grandiosa de mi vida, a ellos les agradezco no haberse rendido y siempre darme la fuerza para seguir adelante, por confiar en mí cuando nadie más lo hacía.

David

“¿Qué significa el desarrollo de la conciencia? Significa algo más profundo que el aprendizaje de teorías estrictamente en los libros; teoría y práctica, ejercicio de la teoría, deben ir siempre unidos, no pueden separarse de ninguna manera, de tal forma que el desarrollo de la conciencia debe estar estrechamente ligado al estudio, al estudio de los fenómenos sociales y económicos que dirigen esta época y a la acción revolucionaria, acción revolucionaria que se traduce en los estudiantes fundamental y primordialmente en estudiar, estudiar, porque esa es la función fundamental del revolucionario colocado en la situación de alumno universitario”

-Ernesto “Che” Guevara-

A mis padres que a pesar de los momentos difíciles que pasamos, con su sacrificio son los principales promotores de mis sueños, por confiar y creer en mí aun cuando pensaba que sería muy tarde para continuar, su tozudez, principios y consejos son las herramientas necesarias para cumplir mis objetivos.

A mis hermanos, personas que iluminan mi vida, con su apoyo alcanzo de mejor manera mis metas, por estar siempre presentes, acompañándome y por el apoyo moral, que me brindaron a lo largo de esta etapa de mi vida.

No puedo dejar de agradecer a Paola, mi compañera fiel de corazón y vida, que, con su amor y respaldo, me ayuda alcanzar mis sueños.

A todos mis amigos, gracias por su amistad incondicional y por compartir algunos de los mejores momentos de mi vida, a mi compañero de tesis David, quien con su esfuerzo y compromiso hemos culminado este proyecto.

A la Universidad Politécnica Salesiana por permitirme concluir con una etapa de mi vida, y de forma especial dar un agradecimiento a mi tutor Ingeniero Gustavo Caiza, gracias por su paciencia, por brindarme sus conocimientos, y por guiarme académicamente con su profesionalismo en el desarrollo de este proyecto.

Y a todos aquellos que de una u otra manera estuvieron involucrados en la culminación este trabajo.

Ernesto

ÍNDICE DE CONTENIDO

CESIÓN DE DERECHOS DE AUTOR	ii
DECLARATORIA DE COAUTORÍA DEL DOCENTE TUTOR	iii
DEDICATORIAS.....	iv
AGRADECIMIENTOS	v
ÍNDICE DE CONTENIDO	vii
ÍNDICE DE TABLAS.....	x
RESUMEN.....	xi
ABSTRACT	xii
INTRODUCCIÓN.....	1
CAPÍTULO 1.....	2
ANTECEDENTES	2
1.1 Planteamiento del problema.....	2
1.2 Justificación.....	3
1.3 Objetivos	4
1.3.1 Objetivo General	4
1.3.2 Objetivos Específicos.....	4
1.4 Tema.....	5
1.5 Beneficiarios de la propuesta	5
CAPÍTULO 2.....	6
ESTADO DEL ARTE	6
2.1 Discapacidad Visual.....	6
2.2 Discapacidad Visual en el Mundo.....	6
2.3 Discapacidad Visual en Latinoamérica	7
2.4 Discapacidad visual en el Ecuador.....	7
2.5 Discapacidad visual y problemas asociados	8
2.6 Métodos educativos para personas con discapacidad visual.....	9
2.7 Dispositivos de lectura	10
2.8 Gafas para personas con ceguera	10
2.9 Raspberry Pi	11
2.10 Python	12
2.11 Librería OpenCV	12
2.12 Reconocimiento Óptico de Caracteres (OCR)	13

2.13	Convertidor TTS (Text To Voice)	13
2.14	Visión artificial.....	14
CAPÍTULO 3.....		15
DESARROLLO.....		15
3.1	Diagrama general de funcionamiento	15
3.2	Diagrama general del prototipo.....	16
3.3	Etapas del sistema	18
3.4	Raspberry Pi 3 B+	19
3.5	RPi UPS PowerPack	19
3.6	Raspberry pi Cámara.....	20
3.7	Método de Visión Artificial	20
3.7.1	GaussianBlur	20
3.7.2	Numpy.....	21
3.7.3	Canny	22
3.7.4	Círculo de cerramiento mínimo	23
3.8	Método de Texto a Voz.....	23
CAPÍTULO 4.....		26
PRUEBAS Y RESULTADOS		26
4.1	Detección de Márgenes	26
4.2	Sistema de detección de caracteres (Tesseract OCR)	28
4.3	Validación del Prototipo	30
4.6.2	Porcentaje de validación del prototipo	31
CONCLUSIONES.....		34
RECOMENDACIONES.....		36
REFERENCIAS		37

ÍNDICE DE FIGURAS

Figura 2.1 Discapacidad visual Grupos etarios.....	8
Figura 3.1 Diagrama General de Componentes.	15
Figura 3.2 Diagrama de flujo de funcionamiento.	17
Figura 3.3 Etapas de las gafas para discapacitados.....	18
Figura 3.4 Imagen Gaussina obtenida con la cámara.....	21
Figura 3.5 Modelo HSV para reconocimiento de colores.....	22
Figura 3.6 Reconocimiento del color verde y dibujo de sus contornos.	23
Figura 3.7 Imagen procesada lista para aplicar Tesseract OCR.....	24
Figura 3.8 Texto resultante de aplicar Tesseract OCR.	25
Figura 4.1 Diagrama de bloques del funcionamiento	26
Figura 4.2 Detección y medición de radio del margen	27
Figura 4.3 Imágenes tomadas por Pi Cámara V2.1	28
Figura 4.4 Resultado de Tesseract OCR	29
Figura 4.5 Prototipo de gafas de lectura para personas con discapacidad visual.....	30
Figura 4.6 Prototipo de gafas para personas con ceguera	31
Figura 4.7 Resultados de validación del prototipo.....	33

ÍNDICE DE TABLAS

Tabla 2.1 Número de personas con discapacidad.	8
Tabla 2.2 Características técnicas de la raspberry Pi 3B+	12
Tabla 3.1 Conexiones para el prototipo	19
Tabla 4.1 Medición de la distancia de reconocimiento del margen.	27
Tabla 4.2 Porcentajes de eficiencia del sistema.	29
Tabla 4.3 Factores de validación del prototipo	32

RESUMEN

Las gafas para lectura de texto con visión artificial permiten mejorar la capacidad de desenvolverse en el entorno a las personas con discapacidad visual, accediendo a todo tipo de información escrita mediante visión artificial y obteniendo una adecuada inclusión de estas personas en las actividades cotidianas. Para la primera parte se determinó la utilización de OpenCV; además consta de una cámara digital la cual toma una foto del documento que se desea leer, luego la imagen pasa a un proceso de mejoramiento, el cual mediante programación en lenguaje Python procesa la misma utilizando técnicas de visión artificial para posteriormente ser implementada en el convertidor de texto a voz. Se realizaron pruebas de funcionamiento para la validación del prototipo con personas de la catedra UNESCO de la Universidad Politécnica Salesiana, para que las gafas faciliten su inclusión. Las pruebas realizadas tomaron en cuenta la distancia a la que el texto es reconocido y la claridad con la que las palabras son pronunciadas por la voz del sistema, en estas dos pruebas se obtuvo un mínimo de distancia de reconocimiento de 20cm y para todos los usuarios entrevistados la claridad, el volumen y la velocidad con el que se escucha el texto es más que adecuado para entender los contenidos. Las personas entrevistadas dan como positivo el sistema como parte complementaria de la enseñanza ya que permite entender mejor los textos además el hardware propuesto no supo ser un inconveniente para su utilización.

ABSTRACT

Eyeglasses for reading text with artificial vision improve the ability of visually impaired people to function in their environment, accessing all types of information written using artificial vision and obtaining an adequate inclusion of these people in daily activities. For the first part the use of OpenCV was determined; in addition, it consists of a digital camera which takes a photo of the document that is wanted to read, then the image passes to a process of improvement, which by means of programming in language Python processes the same one using techniques of artificial vision to be later implemented in the converter of text to voice. Function tests were carried out to validate the prototype with people from the UNESCO Chair of the Salesian Polytechnic University, so that the glasses facilitate their inclusion. The tests performed took into account the distance at which the text is recognized and the clarity with which the words are pronounced by the voice of the system, in these two tests was obtained a minimum recognition distance of 20cm and for all users interviewed the clarity, volume and speed with which you hear the text is more than adequate to understand the contents. The people interviewed consider the system to be positive as a complementary part of the teaching since it allows for a better understanding of the texts and the proposed hardware was not a disadvantage for its use.

INTRODUCCIÓN

En el proyecto mostrado a continuación se desarrolla el diseño e implementación de un sistema con visión artificial para el reconocimiento de texto, y reproducción de audio sintético adaptado a las necesidades de personas con discapacidad visual. Y así puedan acceder a toda la información que se encuentra en documentos impresos de uso cotidiano.

En el mercado existen diferentes prototipos y nuevos desarrollos que son costosos, pero la mayoría no se adaptan a las necesidades específicas que tienen los discapacitados en el país, un producto sin muchos detalles y de bajo costo que sean en beneficio de la persona con discapacidad visual, eliminando la dificultad de obtenerlo por el alto costo. De esta forma se hace uso de la tarjeta de desarrollo Raspberry pi 3 modelo B, el cual cuenta con características de hardware: puertos USB, tarjeta de audio y video, 40 pines GPIO (Entrada/Salida de Propósito General) configurables y de software sistema abierto Raspbian con la librería de OpenCv con sus diferentes utilidades.

Para el desarrollo de las gafas para lectura de texto se desarrollan cinco capítulos: En el capítulo uno se detalló el planteamiento del problema, justificación de la realización de las gafas para lectura de texto, objetivo general y específicos. En el capítulo dos se recopilaron los elementos teóricos necesarios para establecer los fundamentos de la investigación en donde se analizó cada situación de la educación existente que tienen las personas con discapacidad visual, además una revisión a los prototipos existentes, y los diferentes componentes de hardware y software utilizados en el sistema de las gafas para lectura de texto. En el capítulo tres se abordó la selección de los materiales y dispositivos a utilizar en el presente proyecto, desde los elementos a ser utilizados y software con las librerías, funciones y métodos usados. En el capítulo cuatro se realizó el análisis del funcionamiento de todos los sistemas al ejecutar pruebas de operatividad, así como la validación del prototipo con personas de la catedra UNESCO de la Universidad Politécnica Salesiana con algún tipo de discapacidad visual. Por último, el capítulo cinco se muestran las conclusiones y recomendaciones que se dan sobre el proyecto desarrollado. Finalmente, el prototipo tuvo un alto porcentaje de aceptación en su funcionalidad, debido a factores como costo portabilidad y facilidad de uso.

CAPÍTULO 1

ANTECEDENTES

1.1 Planteamiento del problema

La colectividad actual por medio del uso de la tecnología se encargó de reiterar que la información debe ser de fácil acceso, manipulación y reproducción, pero las personas con discapacidad visual, aunque existe, no cuentan con esta tecnología, ni prototipos de bajo costo que les permita tener este tipo de asistencia y apoyo como complemento para el pleno desarrollo de sus capacidades. Además, se suma la dificultad en el manejo del sistema braille en centros educativos por parte de los docentes y estudiantes, lo que conlleva a una baja comprensión del texto, una velocidad de lectura más lenta y un alto costo en la traducción de textos y materiales de estudio necesarios para las personas no videntes.

Según datos y cifras de la Organización Mundial de la Salud refiriéndose a la discapacidad visual, la cifra estimada de personas que viven con alguna forma de deficiencia visual a nivel mundial es de aproximadamente 1300 millones, de estas personas 217 millones tienen una deficiencia visual de moderada a grave y 36 millones son ciegos. La mayoría de las personas con visión deficiente tienen más de 50 años (Organización mundial de la salud, 2017), las principales causas de las enfermedades oculares crónicas que provocan la pérdida de la visión son los errores de refracción y cataratas no operadas.

El Consejo Nacional para la igualdad de Discapacidades (CONADIS) identifica a 461.687 persona con algún tipo de discapacidad registradas hasta mayo de 2019, de las cuales el 11,81% tienen discapacidad visual (Consejo Nacional para igualdad de discapacidades, 2019). Aun cuando unos de los objetivos del Ministerio de Inclusión Económica y Social (MIES) es incrementar el acceso y calidad de los servicios de inclusión social con énfasis en los grupos de atención prioritaria y la población que se encuentra en pobreza o vulnerabilidad, para reducir las brechas existentes; la falta de texto en braille para el estudio, capacitación de profesionales para la enseñanza de braille, y los altos costos que involucra la tecnología (gafas para no videntes) es uno de los inconvenientes en escuelas y colegios regulares. Cuando el niño, joven o adulto con discapacidad visual no tiene este tipo de accesos es por lo que no se puede hablar

de una educación inclusiva. La estrategia de acción para atender a las personas con discapacidad visual debería incluir no solo métodos de aprendizaje para no videntes como el sistema braille o software audio-lector, sino también en el desarrollo de tecnología enfocado en las necesidades específicas de cada grupo de discapacitados visuales en nuestro país, abaratando costos y orientada a la integración e inclusión educativa, social y laboral.

En el ámbito de las TIC, la accesibilidad apunta al conjunto de propiedades que debe incorporar un producto, servicio o sistema, de forma que el mayor número posible de personas, y en el mayor número posible de circunstancias, puedan acceder a él y usarlo (Ruiz, 2017)

Actualmente el acceso a la educación es una de las principales necesidades de los no videntes, porque que les permitirá encontrar un trabajo y que se valore sus conocimientos, más allá de su limitación. El no poder hacerlo los enfrenta a numerosas dificultades.

Actualmente las gafas para discapacidad visual se pueden convertir en una herramienta de gran ayuda para las personas con discapacidad visual, debido a que les permite incluirse en la educación para desarrollar mejor sus habilidades. La realización del prototipo se enfoca en el desarrollo de unas gafas para lectura de texto funcional, dotado de un sistema embebido de bajo costo, para la ayuda de lectura de texto, a un bajo costo y al alcance de las personas con discapacidad visual que en su mayoría son de recursos limitados.

1.2 Justificación

El realizar dispositivos para las personas con discapacidad visual, parte de la necesidad que tienen para desenvolverse fácilmente en ámbitos como la lectura de texto, ser una persona independiente, autónoma, que aporte a la sociedad y a su vez, a no ser discriminada por su discapacidad.

Las TIC (Tecnologías de la Información y la Comunicación) son consideradas herramientas al alcance de un gran número de personas, disminuyendo así las barreras de accesibilidad, en especial de personas con algún grado de discapacidad, permitiendo

una mejor calidad de vida y mejorando su personalidad al sentirse más integrados y participantes activos de la sociedad. (Ruiz, 2017)

Teniendo en cuenta lo antes mencionado, para que se produzca una verdadera integración de los discapacitados visuales a la sociedad, deben contar con una mayor variedad de dispositivos tecnológicos en precios y especificaciones, porque como cualquier ciudadano, ellos tienen funciones sociales a partir del desarrollo de sus aptitudes y capacidades.

Con respecto a la identificación de textos sin la utilización del sistema Braille el audio se constituye en un elemento indispensable para facilitar el proceso de transmisión de palabras que han sido identificadas a partir de un elemento visual tiflotécnicos escogido como el adecuado para facilitar el proceso de aprendizaje en los casos educativos y de inclusión en general para las personas con discapacidad visual.

1.3 Objetivos

1.3.1 Objetivo General

Desarrollar un prototipo de gafas para lectura de texto mediante visión artificial asistido por audio que apoye a personas con discapacidad visual.

1.3.2 Objetivos Específicos

- Recopilar información de papers, documentos y/o prototipos de gafas de lectura para la obtención de las características de desarrollo.
- Diseñar un prototipo de gafas con visión artificial controlado por un sistema embebido para la asistencia de personas con discapacidad visual.
- Implementar un sistema de visión artificial para el reconocimiento de texto en base a la información recopilada.
- Implementar un convertidor de texto a voz compatible con el sistema de visión artificial para la identificación de palabras.
- Realizar las pruebas de funcionamiento para la validación del prototipo con personas de la cátedra UNESCO de la Universidad Politécnica Salesiana.

1.4 Tema

Desarrollo de un prototipo de gafas para lectura de texto con visión artificial que asista a personas con discapacidad visual.

1.5 Beneficiarios de la propuesta

Las personas con discapacidad visual serán beneficiadas en su búsqueda de aprender o escuchar un texto a través de las gafas inteligentes, lo que les permitirá tener mayor confianza al momento de realizar sus estudios, generando la oportunidad de conseguir trabajo logrando finalmente una mayor inclusión y participación en la sociedad. De igual forma el alcanzar un costo accesible permitirá llegar a un mayor número de personas con discapacidad visual podrá mejorar y facilitar su vida.

CAPÍTULO 2

ESTADO DEL ARTE

2.1 Discapacidad Visual

La vista es un sentido global que permite identificar a que distancia y al mismo tiempo que objeto se encuentra frente a nosotros. Las personas con discapacidad visual deben utilizar todos sus otros sentidos para identificar los objetos y personas: el gusto, el tacto, el oído y el olfato, les permiten percibir el mundo a su alrededor.

Según la Organización Mundial de la Salud (OMS) la discapacidad visual, “puede limitar a las personas en la realización de tareas cotidianas y afectar su calidad de vida, así como sus posibilidades de interacción con el mundo circundante” (Organizacion mundial de la salud, 2017).

2.2 Discapacidad Visual en el Mundo

Según datos de la OMS hasta octubre de 2018 menciona en su página web, que existe en el mundo alrededor de 1300 millones de personas que viven con alguna forma de deficiencia visual, de las cuales 188,5 millones de personas tienen una deficiencia visual moderada, 217 millones tienen una deficiencia visual de moderada a grave y 36 millones son ciegas (Organizacion mundial de la salud, 2017). Se estima que el 90 % de personas en el mundo con discapacidad visual provienen de países en desarrollo. Las causas principales de pérdida de la visión son enfermedades oculares, como, por ejemplo: errores de refracción, cataratas no corregidas y operadas.

El Consejo Nacional de Fomento Educativo (CONAFE) de la Secretaría de Educación Pública de México define la agudeza visual como la capacidad que tiene una persona para percibir con claridad y nitidez, formas y figuras a determinada distancia. Las personas con agudeza visual normal registran una visión de 20/20. La gente que usa lentes se las considera que tiene un problema de agudeza visual, pero no se consideran personas con discapacidad visual ya que esta deficiencia puede ser sustentada con gafas graduadas.

Según la OMS la ceguera es considerada como pérdida de visión en cualquiera de los dos ojos, al momento de tener la visión por debajo de la agudeza visual de 20/200, aun

después de realizar correcciones mediante tratamientos oftalmológicos y el uso de lentes de contacto o gafas. (Organizacion mundial de la salud, 2017)

La agudeza se mide con dos números. El primer número, 20 (6 metros), es la distancia de la persona al objeto mostrado. El segundo número, es la distancia donde la persona con una vista "perfecta" o "normal" estaría situada para poder ver de la misma forma que la primera.

Lo que se puede deducir en lo mencionado anteriormente es que la discapacidad visual puede ser el resultado de muchos factores adversos, es decir tener un contexto ambiental desfavorable y no depende únicamente de las características físicas o biológicas de la persona. Por tal motivo, las personas con discapacidad visual siempre han buscado medios y modos alternativos que les permita tener un mejor acceso y movilidad, para que este tipo de discapacidad no los aleje de la vinculación con la sociedad ni simbolice más segregación.

2.3 Discapacidad Visual en Latinoamérica

En publicaciones y estudios muestran a las cataratas como la principal causa de ceguera en la región. A nivel mundial, la catarata, el tracoma, la oncocercosis, la ceguera infantil, y los errores refractivos son las prioridades primarias de la Organización Mundial de la Salud (Organizacion mundial de la salud, 2017).

En lo que respecta a los países latinoamericanos la ceguera afecta entre el 1% y 4% a esta parte de la población, esto representa una cantidad importante para estos países, considerando que están en desarrollo y que sufren un gran impacto en la economía al no contar con trabajo entre el 2 al 8% de la población. Cuando hablamos del total de personas ciegas, la patología de mayor prevalencia es la Catarata que afecta a aproximadamente 60%, seguida de la retinopatía diabética y el glaucoma, principalmente. (VISION 2020, 2014).

2.4 Discapacidad visual en el Ecuador

En Ecuador según el Consejo Nacional para la igualdad de discapacidades (CONADIS) existen un total de 54529 personas registradas con discapacidad visual hasta junio de 2019, lo que corresponde 11.87% del total de personas con algún tipo

de discapacidad en el Ecuador, ya sea moderada o grave, entre otros casos (Consejo Nacional para igualdad de discapacidades, 2019).

En la Tabla 2.1 se observa el porcentaje de acuerdo al grado de discapacidad visual registradas a nivel nacional.

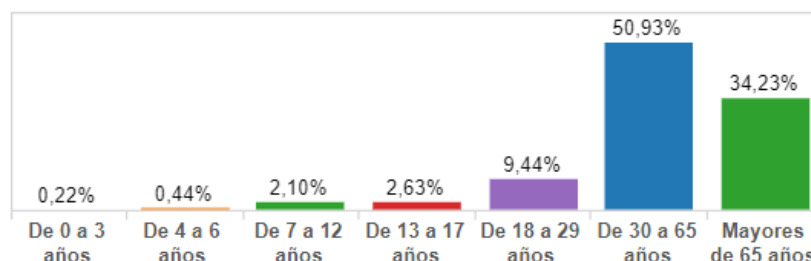
Tabla 2.1 Número de personas con discapacidad.

Grado de discapacidad visual	Numero personas con discapacidad
30% - 49 %	18481
50% - 74%	14642
75% - 84%	16217
85% - 100%	5189
Total	54529

Personas con discapacidad visual Registradas Fuente: (Consejo Nacional para igualdad de discapacidades, 2019)

En la Figura 2.1, se representa de forma porcentual la información de grupos etarios, donde se observa que el mayor número de discapacitados visuales registrados se encuentran en la edad de 30 a 65 años.

Figura 2.1 Discapacidad visual Grupos etarios.



Representación en forma porcentual de los grupos etarios con discapacidad visual en el Ecuador Fuente: (Consejo Nacional para igualdad de discapacidades, 2019)

2.5 Discapacidad visual y problemas asociados

El Gobierno Nacional con el apoyo de Organismos Internacionales ha instaurado y fortalecido las instituciones encargadas de implementar proyectos y programas que garanticen el derecho de las personas con discapacidades a acceder a salud, educación y empleo. Sin embargo, aún existe un gran porcentaje que tienen la necesidad de estudiar y obtener plazas de trabajo dignas que vayan acorde a sus capacidades,

habilidades y conocimientos. La población con discapacidad visual en pichincha es de 8592, siendo tan solo 104 los estudiantes con discapacidad en educación básica media y bachillerato registrados, de acuerdo a las estadísticas disponibles generadas por el Instituto Nacional de Estadísticas y Censos (INEC, 2019).

La educación es un proceso de constante aprendizaje cultural, social, de doctrinas y normas morales de la sociedad. Sin este aspecto a una persona le sería muy difícil el desarrollo personal o la inserción activa y consciente en el mundo social.

Los alumnos con discapacidad visual (sin otra discapacidad) logran integrarse al aula regular y realizar la mayoría de las actividades junto con el resto de sus compañeros del grupo, siempre y cuando se le ofrezcan apoyos específicos. Los apoyos abarcan materiales específicos, personas y estrategias metodológicas y de intervención que el docente ofrece a los alumnos con discapacidad visual para que alcancen los objetivos propuestos en el grupo en el que se encuentran integrados. (CONAFE, 2016).

2.6 Métodos educativos para personas con discapacidad visual

Las TIC (Tecnologías de la Información y Comunicación) y las tecnologías adaptativas son en la actualidad el campo más significativo en lo que a innovar y desarrollar se refiere, convirtiéndose en el recurso más importante para que las personas con discapacidad visual puedan acceder a la información y realizar tareas de lectura y escritura en forma autónoma.

Muchos programas informáticos tratan de compensar las limitaciones educativas que tiene una persona con discapacidad visual, en programas tradicionales como los lectores de formatos PDF (Portable Document Format) se crean con el objetivo de leer un archivo, la cual toma el texto del archivo y la convierte en audio.

Además, en los softwares informáticos, existen métodos y sistemas que permiten a la persona invidente acceder a la información como es el método Braille. La tecnología NVDA (NonVisual Desktop Access) realiza una traducción de los textos ingresados en una portátil en voz, así cualquier información que ingrese podrá ser escuchada.

2.7 Dispositivos de lectura

La mayor barrera para los discapacitados visuales a la hora de buscar información es el manejo de texto, siendo los dispositivos de lectura los que están especialmente diseñados para las personas con discapacidad visual.

Existen dispositivos que sirven para leer documentos como el desarrollado por los Investigadores del MIT Media Labs, un dispositivo llamado anillo lector. Según se lee en Tecnobility, el anillo está conectado a un ordenador que traduce el texto escrito y se lo lee al usuario, permitiendo que personas que tengan pérdida de visión puedan leer un libro, un periódico, una revista o cualquier texto escrito. Además, este dispositivo ayuda al usuario en el caso de que se desvíe de la línea de texto o cuando llega al final de la línea, del párrafo o de la página (TECNOBILITY, 2019).

Otro tipo de dispositivo es la solución Ablecenter que permite al usuario con baja visión ver en su pantalla de ordenador a un tamaño mucho más grande cualquier objeto, imagen o texto que haya en la habitación, sin necesidad de utilizar dispositivos adicionales como lupas. De esta forma, con solo conectar un dispositivo en la parte superior de la habitación, ese lugar se convertirá en un espacio accesible para cualquier persona con discapacidad visual, facilitando que pueda seguir el desarrollo de reuniones o conferencias sin otros aparatos adicionales (TECNOBILITY, 2019).

2.8 Gafas para personas con ceguera

Actualmente se están desarrollando diferentes tipos de gafas con tecnologías que utilizan cámaras con 3 D, reconocimiento de objetos, etc. Y que, mediante diferentes programas de software, proporciona ayuda al discapacitado visual.

Los docentes del área de Investigación de la Universidad Carlos III de Madrid (UC3M) desarrollaron un sistema que integra unas gafas de realidad virtual que ayudan a las personas con discapacidad visual a detectar la distancia y forma a la que se encuentran los objetos.

Las lentes de Eyesynth, según Diario de Ibiza, tiene como objetivo el de conseguir una "experiencia de sentido aumentado". O lo que es lo mismo: que los ciegos puedan 'ver' a través del oído. Lo primero que hay que dejar claro es que estas gafas inteligentes no

permiten ver a las personas ciegas (Arráez, 2017). Generalmente son de lectura de texto o que ayudan a orientarse espacialmente a las personas con discapacidad visual.

En este caso las cámaras situadas en la montura de las gafas escanean y registran el espacio creando imágenes en tres dimensiones. Toda esa información es procesada por el ordenador que hay en el interior de la petaca y posteriormente es convertida en señales sonoras que se transmiten a través de unos auriculares cocleares. Tras un poco de entrenamiento, el usuario es capaz de detectar e identificar objetos a su alrededor, tanto fijos como móviles. (Arráez, 2017)

El presente proyecto busca diseñar y construir las gafas para discapacitados visuales capaz de leer documentos impresos, mediante un proceso intuitivo y fácil aprendizaje.

2.9 Raspberry Pi

Es un pequeño ordenador de placa reducida fabricado en Reino Unido. La primera versión fue lanzada al mercado en el año 2012 con el propósito de inspirar la enseñanza de ciencias de computación. La tarjeta tiene pines GPIO (General Purpose Input/Output) de entrada y salida de propósito general y mediante estos se tiene una interfaz con el mundo exterior. Estos pines son una interfaz física. Tiene 40 pines de los cuales 26 son GPIO y los otros son de alimentación de 3,3 V, 5 V y tierra, además, dos pines de ID EEPROM (Caiza & Garcia , 2017). En el presente proyecto se va a utilizar la tarjeta Raspberry Pi 3B+ debido a sus altas y mejoradas prestaciones respecto a los modelos anteriores, en la Tabla 2.2 se puede observar algunas de sus características principales mencionadas.

Cuentan con un chip Broadcom BCM2835, el cual posee un procesador ARM11 con diferentes frecuencias de trabajo, un procesador gráfico Video Core, además, presentan una salida de audio y video por medio de un conector HDMI que permite conexión con monitores o televisiones para conexiones de este tipo. (Rodrigo, 2014)

Tabla 2.2 Características técnicas de la Raspberry Pi 3B+

	Raspberry Pi 3B+
CPU	Broadcom BCM2837 de 64 bits
Velocidad del procesador	QUAD Core @1.2 GHz
RAM	1GB SDRAM@1250
USB 2.0	4x Puertos USB
GPIO	Extendido de 40 40 pines
Chipset del procesador	Broadcom BCM2837 64Bit ARMv7 Procesador Quad Core Single Board Computer running at 1250MHz
Almacenamiento	Puerto Micro SD para cargar su sistema operativo y almacenar datos
Consumo de energía/voltaje	Fuente de alimentación Micro USB conmutada mejorada hasta 2.5 ^a

Características, datos de almacenamiento y operación de la Raspberry Pi 3B, Fuente: (RO-BOTICA, 2019)

2.10 Python

Python es un lenguaje de programación multipropósito de alto nivel. Soporta orientación a objetos, programación imperativa y, en menor medida, programación funcional. Es un lenguaje interpretado, usa tipado dinámico y es multiplataforma. Su filosofía de diseño enfatiza la productividad del programador y la legibilidad del código.

Python es un lenguaje de programación poderoso y fácil de aprender. Cuenta con estructuras de datos eficientes y de alto nivel y un enfoque simple pero efectivo a la programación orientada a objetos. La elegante y limpia sintaxis de Python y su tipado dinámico, junto con su naturaleza interpretada, hacen de éste un lenguaje ideal para scripting y desarrollo rápido de aplicaciones en diversas áreas y sobre la mayoría de las plataformas (Rodrigo, 2014).

2.11 Librería OpenCV

OpenCV (Open Source Computer Vision Library) es una biblioteca abierta de visión artificial, desarrolladas por Intel. Su primera aparición fue en 1999 en el cual presentaron su primera versión alfa. Esta librería cuenta con diferentes funciones para el procesamiento de imágenes, esto es debido a su licencia BSD, la cual permite que

se utilice de manera libre para investigaciones y propósitos comerciales en una inmensidad de aplicaciones (Rubio, 2015).

Uno de los objetivos de OpenCV es proveer una infraestructura de visión por computador fácil de utilizar que ayuda a los programadores a desarrollar aplicaciones ‘sophisticadas’ de CV (Computer Vision) rápidamente. La librería OpenCV contiene aproximadamente 500 funciones que abarcan muchas áreas de CV, incluyendo inspección de productos de fábricas, escaneo médico, seguridad, interfaces de usuario, calibración de cámaras, robótica...etc, porque la visión por computador y el aprendizaje automático van de la mano. OpenCV también tiene una completa librería de uso general de aprendizaje automático (MLL o Machine Learning Library), la cual es muy útil para cualquier problema de aprendizaje automático. Esta sublibrería está especializada en el reconocimiento estadístico de patrones y clustering (OSL.ULL, 2015)

2.12 Reconocimiento Óptico de Caracteres (OCR)

Permite ver texto es decir es una herramienta que lee texto de imágenes y fotografías. Se realiza una captura de texto o documentos e incluso de señales en las calles.

El OCR (Optical Character Recognition) es una herramienta tecnológica que permite emular la capacidad del ojo humano para reconocer caracteres. Para que este software reconozca cada caracter, inspecciona la imagen pixel a pixel, buscando coincidencias con las fuentes que usamos normalmente. El nivel de precisión de la herramienta OCR es bastante fiable, pero debemos recordar que, aunque se tenga una base de fuentes muy amplia para comparar las imágenes, el sistema no es cien por ciento efectivo, por ello se debe revisar el texto obtenido para evitar tener malas traducciones o frases sin sentido.

2.13 Convertidor TTS (Text To Voice)

Este tipo de convertidor es usado para tomar un archivo de texto, en este caso texto plano (.txt), analizarlo y procesarlo para convertir su información en audio para que pueda ser reproducido por voz artificial.

2.14 Visión artificial

La Visión Artificial o Visión por Computador es un área multidisciplinar que pretende, en cierta medida, reproducir artificialmente el sentido de la vista mediante el procesamiento e interpretación de imágenes, captadas con distintos tipos de sensores (fundamentalmente, cámaras), y utilizando para ello las prestaciones de los ordenadores. El desarrollo de sistemas de Visión Artificial requiere una combinación de etapas de más bajo nivel (p. ej. operaciones de eliminación de ruido o de aumento del contraste) para mejorar la calidad de las imágenes capturadas junto con otras etapas de más alto nivel (de reconocimiento de patrones y de interpretación de imágenes) para reconocer los elementos presentes en una escena. Las aplicaciones de la Visión Artificial son, en general, muy dependientes del problema concreto a resolver. A modo de ejemplo, pueden destacarse las relacionadas con la medicina, la inspección automática, la geología, el análisis de documentos, etc. (Sanchez, 2005)

CAPÍTULO 3

DESARROLLO

En el presente capítulo se detalla todo lo que corresponde al prototipo de gafas para personas con discapacidad visual, su funcionamiento y construcción, como responde cada elemento de hardware y software en todo el sistema.

3.1 Diagrama general de funcionamiento

El prototipo tiene como base general la incorporación de una cámara (PiCámara) adherida a unas gafas a través de un modelo impreso diseñado específicamente para esta cámara, los lentes son oscuros, como los que normalmente usan las personas con discapacidad visual, esta cámara es la encargada de identificar un marco o margen de color verde dibujado en la hoja con el texto que se desea leer, además el sistema cuenta con un sistema embebido, en este caso la Raspberry Pi3 B+ la cual procesa toda la información adquirida por la PiCámara y la convierte en un audio, al final para escuchar el resultado del texto se usa los audífonos que van conectados al jack de 3.5mm, ya incorporado en la Raspberry, donde se conectan los audífonos. También tiene conectada una batería para Raspberry de LiPo lo que permitirá la portabilidad del prototipo. En la Figura 3.1 se muestra los elementos utilizados para el diseño y construcción del prototipo.

Figura 3.1 Diagrama General de Componentes.

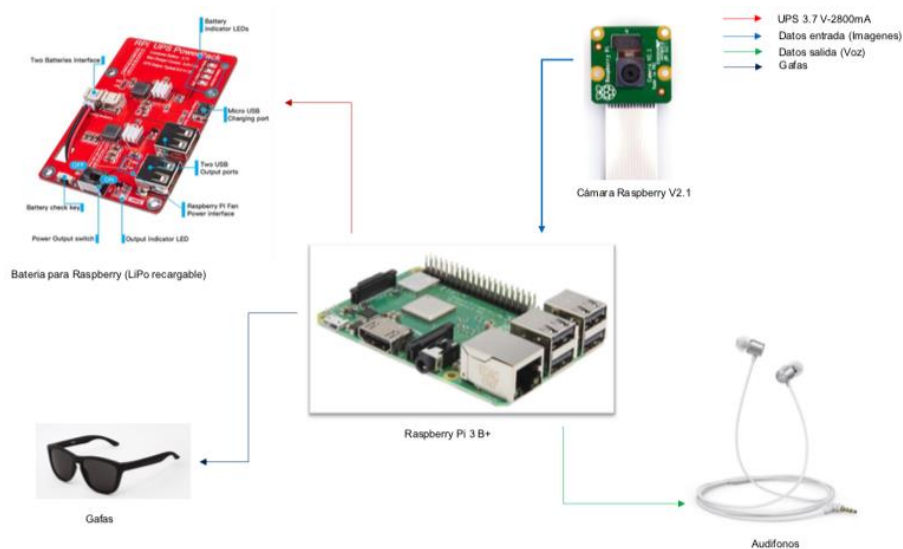


Diagrama general de Componentes de Gafas para lectura de texto.
Elaborado por: David Basantes, Ernesto Chalaco

3.2 Diagrama general del prototipo

En la Figura 3.2 se puede observar el diagrama del prototipo que controla todo el funcionamiento de las gafas, en primera instancia se tiene el encendido del prototipo, suficiente para poner a trabajar todo el sistema, el proceso continúa con una condición de ubicar el texto frente a las gafas moviendo lentamente hacia adelante y atrás, hasta obtener en la salida de audio con el mensaje ENFOCANDO que informa sobre la detención del marco a una distancia correcta. Juntamente con el mensaje ENFOCANDO se debe escuchar el mensaje TOMANDO FOTO que es la advertencia para mantenerse inmóvil durante un segundo, el texto y las gafas, para capturar la imagen con la cámara de la Raspberry Pi, mientras no se escuche los mensajes mencionados el sistema se mantendrá en un bucle esperando dichas confirmaciones. Cuando la imagen ha sido capturada pasa a la instancia de procesamiento de imagen donde se obtiene un archivo (.txt) que contiene el texto que se va a transformar en audio, se debe recalcar que la foto se tomará con una resolución HD, luego el texto obtenido pasa a voz con el convertidor TTS transformando la información en audio para que pueda ser reproducido por voz artificial, finalmente el proceso se repite hasta cuando el usuario apague el sistema.

Figura 3.2 Diagrama de flujo de funcionamiento.

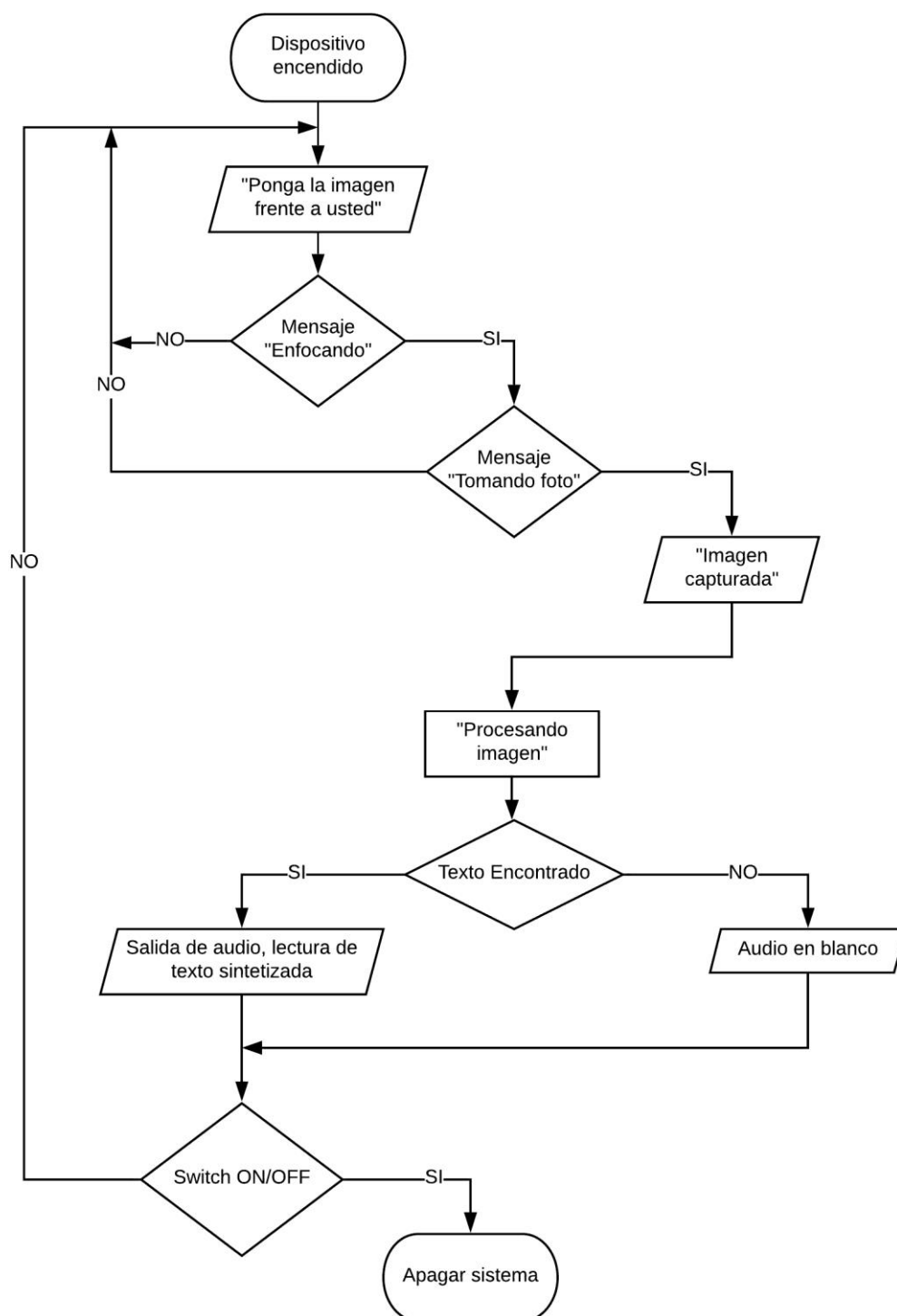


Diagrama de flujo del prototipo de Gafas para lectura de texto.

Elaborado por: David Basantes, Ernesto Chalaco

3.3 Etapas del sistema

Las etapas que tiene las gafas buscan conseguir un proceso de limpieza del texto ante la cantidad de información que se puede adquirir en una captura, el procesamiento analiza el contraste de la imagen en escala de grises para finalmente tomar la decisión de segmentar y reconocer todos los caracteres existentes en la captura de imagen; el proceso se divide en cuatro etapas y se representa en la Figura 3.3

Figura 3.3 Etapas de las gafas para discapacitados



Etapas que intervienen en las gafas para discapacitados. Elaborado por: Ernesto Chalaco, David Basantes.

3.3.1 Captura

Es esta etapa mediante una señal auditiva se establece el momento de captura de la imagen, el mismo que se da cuando el texto sea legible y esté a una distancia correcta, ni muy cerca ni muy lejos, en ese momento la imagen pasa a ser procesada en el sistema embebido, generalmente se utilizan cámaras con una buena resolución con el fin de digitalizar de manera correcta la imagen. Y el proceso se verifica cuando se escucha el mensaje de confirmación FOTO TOMADA.

3.3.2 Procesamiento

Una vez realizada la captura de imagen, se realiza un tratamiento por medio de filtros, esto con el fin de tener las condiciones adecuadas y óptimas para la transformación del texto, para ello se utilizan una variedad de técnicas, como por ejemplo la transformación a escala de grises, además de obtener el gradiente de la imagen con el fin de obtener mayor contraste en los tonos negros para adquirir ciertas propiedades que se necesitan y así evitar errores en la traducción del texto.

3.3.3 Transformación a texto

Esta etapa se centra en conseguir transformar la imagen en un texto plano (.txt) en donde se reconocen todos los caracteres, para posteriormente transformarlo en audio.

la imagen se convierte a texto mediante el reconocimiento óptico de caracteres (OCR) que permite leer texto de imágenes. El tiempo de procesamiento puede tardar unos pocos segundos o varios minutos, esto de acuerdo con la cantidad de palabras que se encuentra en la imagen, entre mayor sea el número de palabras mayor será el tiempo de procesamiento.

3.3.4 Transformación en audio

Finalmente, la última etapa culmina con la reproducción de texto en audio a través de un audiófono (Jack 3.5 mm) o parlante donde se escuchará las palabras encontradas en la imagen capturada al inicio del proceso. Se utiliza el convertidor TTS (Text to Speech) que es una librería del Open CV, convirtiendo la información en audio para que pueda ser reproducido por voz artificial.

3.4 Raspberry Pi 3 B+

Esta tarjeta controladora tiene características que se las ha considerado óptimas para el desarrollo del prototipo y lograr el procesamiento adecuado de la información (Tabla 3.1), entre estas, esta su procesador de 1.4GHz a 64 bits, 1Gb de Ram, que junto con la plataforma Python brinda un soporte amplio para programación orientada a objetos, lo que admite desarrollar algoritmos que hagan prácticamente cualquier cosa, además su hardware es de fácil configuración, así como poseer un tamaño adecuado para transportarlo en la mano.

Tabla 3.1 Conexiones para el prototipo

Raspberry Pi 3B+	Descripción
Conector SCI	Raspberry Pi Cámara V2.1
Conector jack 3.5mm	Audífonos
Conector micro USB	Batería recargable de LiPo, 3.7V

Conexiones Gafas para lectura de texto. Elaborado por: David Basantes, Ernesto Chalaco

3.5 RPi UPS PowerPack

La utilización de esta tarjeta facilita la portabilidad de la batería que permite obtener de forma eficiente el voltaje y la corriente adecuados para el funcionamiento del sistema, es una tarjeta de expansión que tiene dos ranuras para colocar hasta dos batería

de LiPo con un máximo de 5V, este módulo tal como viene proporciona una fuente de alimentación de 3.7V a 3800mA con una corriente máxima nominal de salida de 2A, además de contar con dos puertos USB, uno de ellos conectado a la Raspberry para proporcionarle la energía necesaria para encenderla, finalmente esta tarjeta cuenta con un puerto micro-USB para cargar las baterías y un switch de ON-OFF que servirá para encender y apagar la fuente de alimentación.

3.6 Raspberry pi Cámara

El módulo PiCamera V2.1 viene integrado con un sensor Sony IMX219 de 8mpx, dando una calidad de imagen de 1080p30, ideal para captar el entorno en tiempo real, además de tomar fotos con una resolución de 3280 x 2464 pixeles, permitirá obtener una fotografía de gran tamaño y de calidad excelente, al ser una imagen de alta resolución se podrá detectar de manera adecuada los caracteres en el procesamiento digital de la imagen.

3.7 Método de Visión Artificial

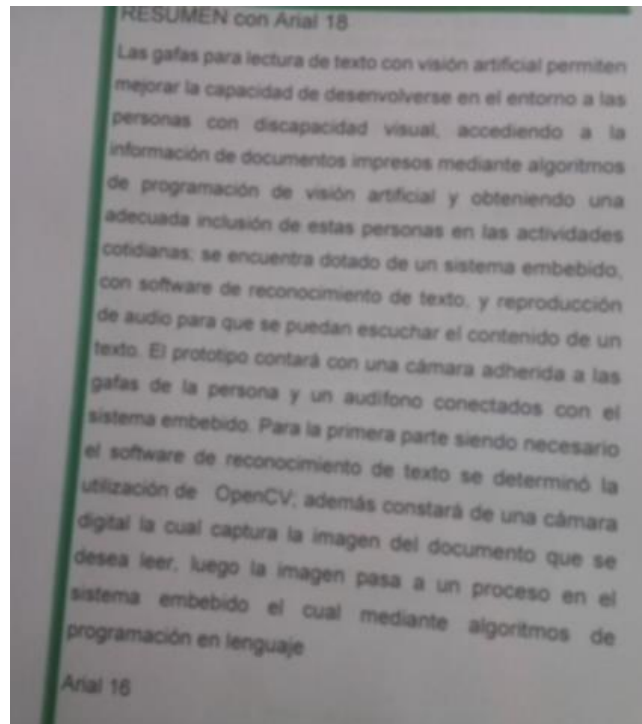
Para obtener los mejores resultados dentro del sistema de visión artificial, es necesario tener instaladas algunas librerías importantes, estas son: OpenCv, esta es la más importante de todas ya que es la que contiene toda la infraestructura en la cual se va a desarrollar el programa, también se debe instalar la librería PiCamera, que, por medio de la interfaz CSI (Camera Serial Interface) se podrá configurar la cámara de acuerdo a los requerimientos de imagen dispuestos para reconocimiento de caracteres, por último se tendrá que obtener la librería Numpy que tiene como objetivo contribuir a los arreglos de imagen, necesarios para su procesamiento.

3.7.1 GaussianBlur

Esta función permite realizar el desenfoque de la imagen además de eliminar ruidos de alta frecuencia en cada pixel. El número de Kernel se especifica como positivo e impar desde 0, considerado en ancho y alto, si se aumenta este valor el desenfoque aumenta proporcionalmente. La desviación estándar se la considera en x, este dato es

el ancho de la campana de Gauss, si es cero, OpenCV calcula el valor de acuerdo al kernel implementado.

Figura 3.4 Imagen Gaussina obtenida con la cámara

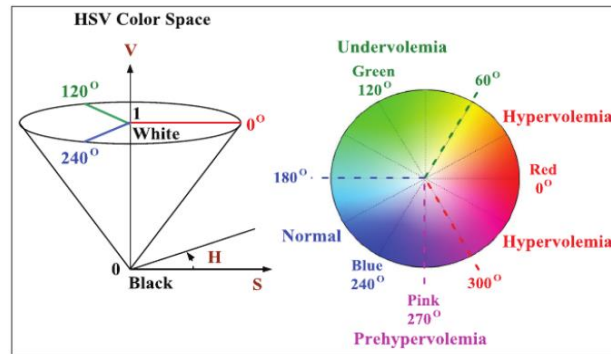


Elaborado por: David Basantes, Ernesto Chalaco

3.7.2 Numpy

Esta librería es necesaria para realizar los arreglos correspondientes al reconocimiento de una imagen por medio de un color específico, para esto cuenta con la función `cv2.COLOR_BGR2HSV`, el cual se usa el método HSV (Hue, Saturation, Value), el matiz o tono va desde 0 a 179, la saturación de 0 a 255 y el valor o brillo de 0 a 255, con este rango de valores se obtiene el color que se desea identificar en la imagen, como se puede ver en la Figura 3.5

Figura 3.5 Modelo HSV para reconocimiento de colores



Modelo HSV para reconocimiento y transformación de colores Elaborado por: (Chen Wei-Ling, 2017)

Se obtuvieron y analizaron los datos para los tonos bajos y altos de color verde, estas variables se almacenan para tener una máscara de color que permita establecer el rango adecuado que detecte el color eficientemente. Obteniendo los siguientes valores:

```
bajos = np.array([27,59,36], dtype=np.uint8)
```

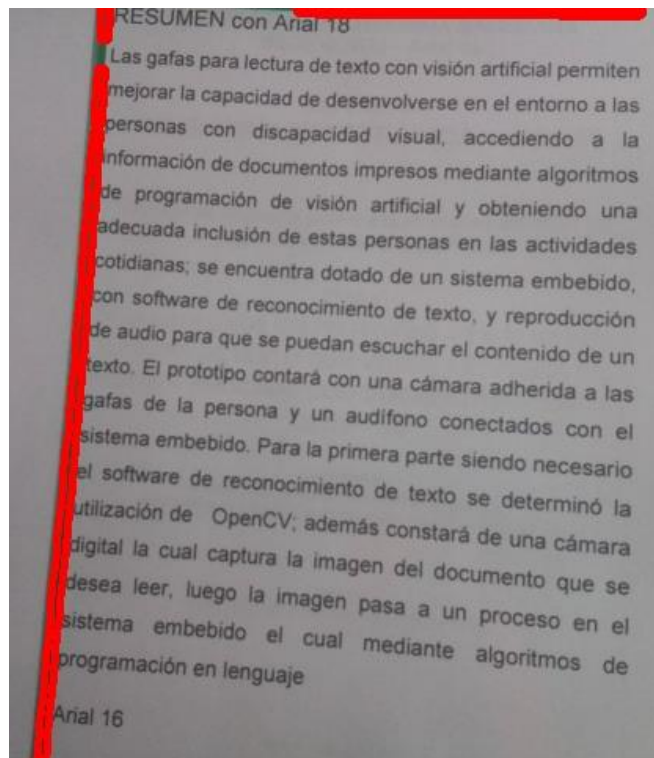
```
altos = np.array([89, 168, 224], dtype=np.uint8)
```

```
mask = cv2.inRange(hsv, bajos, altos)
```

3.7.3 Canny

La función Canny realiza un proceso dividido en tres partes específicas: primero se realiza la detección de bordes por medio de Sobel que consiste en cálculo de la primera derivada para detectar cambios de intensidad, segundo es la supresión de píxeles fuera de borde, se le conoce como non-máximo que es una técnica que permite adelgazar los bordes por medio del gradiente y serán aquellos que tengan un grosor igual a 1, tercero se aplica el umbral por histéresis, permite segmentar una imagen usando un umbral mínimo y otro máximo así se podrá detectar si un píxel forma parte de un borde o no. Al finalizar el reconocimiento de un contorno se hará dos operaciones más: la primera es buscar el contorno por medio de la función `cv2.findContours()`, y finalmente al encontrar el contorno se lo dibuja de la siguiente forma `cv2.drawContours()` como se observa en la Figura 3.6

Figura 3.6 Reconocimiento del color verde y dibujo de sus contornos.



Elaborado por: David Basantes, Ernesto Chalaco

3.7.4 Círculo de cerramiento mínimo

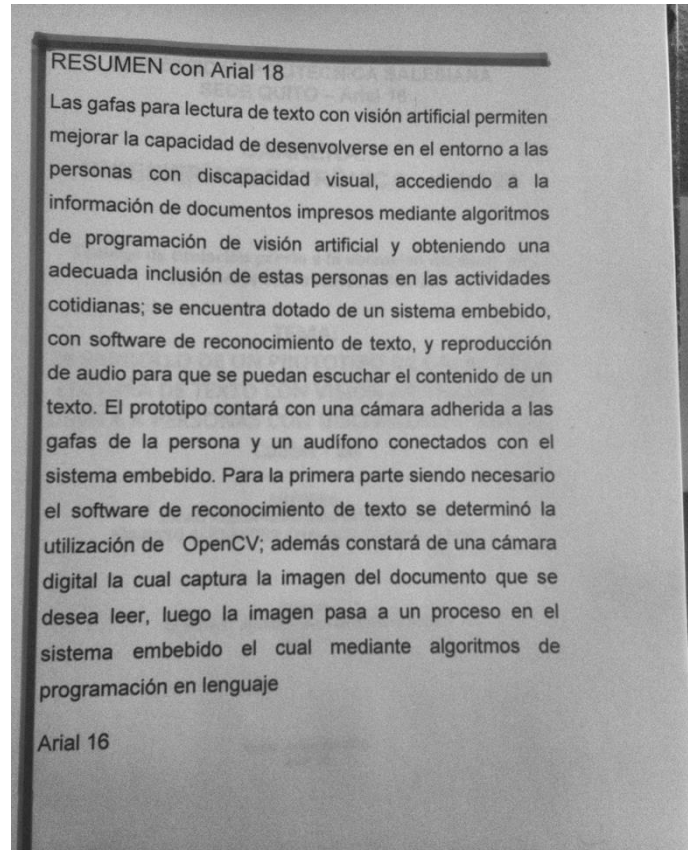
Con esta función que proporciona OpenCv, que se determina como `cv2.minEnclosingCircle()`, se va a encerrar los bordes encontrados en la imagen dentro de un círculo con un radio determinado, de esta forma y con el radio se logra establecer el tamaño del borde encontrado. Instaurando los valores de radio adecuados se logra establecer la distancia a la que la hoja de texto debe ser reconocida, esto sirve para delimitar el espacio de enfoque y evitar que se detecten otros objetos no deseados, ya sea que se encuentren lejos o cerca de la cámara. Al obtener estos datos se pueden usar para establecer la posición en la que se debe tomar la foto de manera intuitiva.

3.8 Método de Texto a Voz

Para procesar la imagen obtenida por la cámara se tiene que previamente haber instalado las librerías de Tesseract OCR y TextToVoice, la primera realiza el proceso en el cual la foto obtenida pasa a transformarse en un texto plano (.txt) en donde se reconocen todos los caracteres, y al final este documento se procesa por medio de

festival voice que se encarga de reproducir los fonemas de cada carácter encontrado en el orden correspondiente, como si alguien estuviera leyendo la hoja.

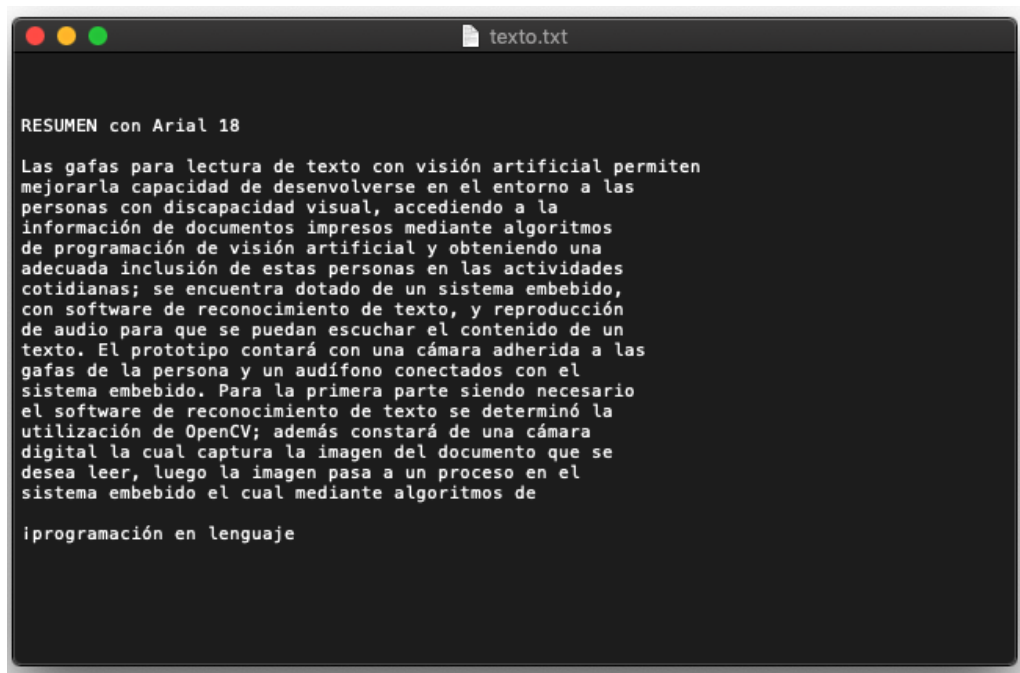
Figura 3.7 Imagen procesada lista para aplicar Tesseract OCR.



Elaborado por: David Basantes, Ernesto Chalaco

Como se observa en la Figura 3.7 la imagen esta lista para empezar el reconocimiento de caracteres por medio de la siguiente función: `os.system('imagen.jpg texto -l spa')`, esta se encarga de transformar esta foto en un texto plano, como se puede ver en la Figura 3.5. Este documento estaría listo para pasar a ser leído, para ello se usa la siguiente sentencia OpenCv, `os.system('less texto.txt | iconv -f utf-8 -t iso-8859-1 -s -c /festival --tts --language spanish')`, ahora se debe esperar a que procese el texto a voz, como se observa en la Figura3.8, este tiempo de espera depende del contenido, es decir, entre más caracteres mayor es el tiempo de procesamiento.

Figura 3.8 Texto resultante de aplicar Tesseract OCR.



Elaborado por: David Basantes, Ernesto Chalaco

CAPÍTULO 4

PRUEBAS Y RESULTADOS

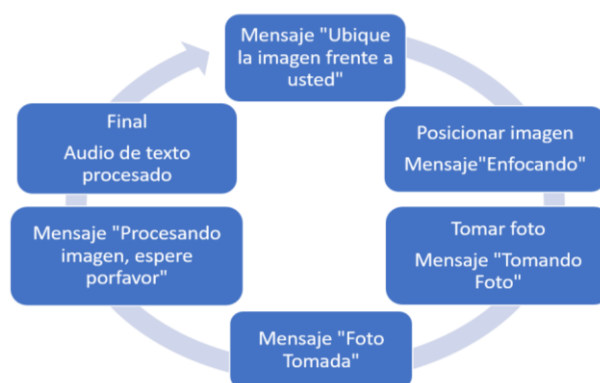
En el capítulo 4 se analiza los resultados del prototipo de gafas para lectura de texto y las pruebas de operabilidad en múltiples situaciones: sistema de posicionamiento y procesamiento de imagen para la traducción a audio del texto.

Para realizar las pruebas se toma en cuenta ciertos parámetros que tienen los documentos que se van a leer, primero hay que dibujar dos líneas verdes de 5mm de espesor como mínimo, una línea en la parte superior o inferior y la otra en la parte izquierda o derecha de la hoja, además el tamaño de letra debe ser de 14 o superior y en letra imprenta. Se decidió colocar el margen rectangular debido a su practicidad al momento de dibujarlo alrededor de un texto y de color verde por la facilidad de detección al utilizar visión artificial, resultando una disminución importante del error durante la detección de este. Con estas indicaciones empiezan las pruebas con diferentes tamaños de letras y contenidos desarrollando los diferentes bloques de funcionamiento como se ve en la Figura 4.1

4.1 Detección de Márgenes

La distancia a la que se puede obtener los mejores resultados está entre 20 y 30cm desde las gafas, donde está ubicada la cámara y la hoja que contiene el texto. El margen rectangular presenta una gran ventaja al desechar cualquier tipo de contornos de diferentes formas y tamaños existentes en la imagen.

10 Figura 4.1 Diagrama de bloques del funcionamiento



Elaborado por: David Basantes, Ernesto Chalaco

Para saber si la hoja está en posición y se ha tomado la fotografía el sistema indica dos estados por medio de una salida de voz, el primero es el ENFOCANDO que determina que la hoja está en posición adecuada para ser capturada. Luego la segunda salida de voz TOMANDO FOTO es cuando la cámara empieza el conteo regresivo para hacer la captura, lo que significa que se debe mantener estática la hoja y las gafas, estas dos alertas se dan gracias al cumplimiento de las condiciones que debe tener en tamaño y radio el margen como se puede apreciar en la Figura 4.2.

Figura 4.2 Detección y medición de radio del margen

```
He encontrado 1 objetos
nmero de radios: 1
frecuencia 2
He encontrado 1 objetos
nmero de radios: 1
frecuencia 3
He encontrado 1 objetos
nmero de radios: 1
frecuencia 4
He encontrado 2 objetos
nmero de radios: 1
frecuencia 1
He encontrado 1 objetos
nmero de radios: 1
frecuencia 2
He encontrado 2 objetos
nmero de radios: 1
frecuencia 1
```

Datos del radio encontrado, Elaborado por: David Basantes, Ernesto Chalaco

Se hizo pruebas con 6 hojas, los resultados se observan en la Figura 4.1, que es la distancia donde mejor enfoque y mayor cantidad de palabras traducidas se obtuvo para cada hoja, la distancia es desde el lente de la cámara hacia la hoja que contiene el texto:

Tabla 4.1 Medición de la distancia de reconocimiento del margen.

Hoja	Distancia [cm]
A	24
B	22
C	26
D	28
E	25
F	21
Promedio	20.8

Distancias a las que se verifica el margen. Elaborado por: Ernesto Chalaco, David Basantes

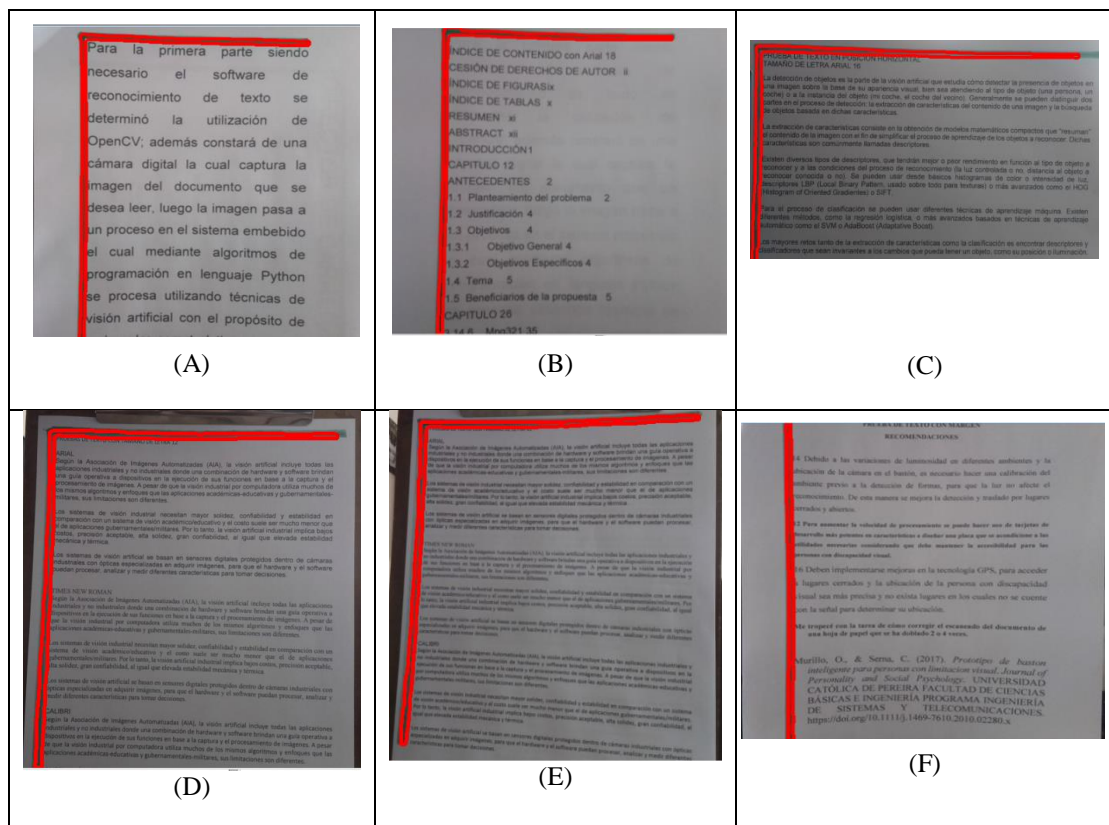
El promedio de distancia entre la cámara y la hoja de texto está en 20,8 cm, y está relacionada directamente con la resolución, enfoque (ajustado manualmente) y campo

visual de la cámara, además de la configuración realizada previamente en la programación de visión artificial. Además de la distancia la calidad de la imagen se puede ver afectada por el tamaño del margen, y por supuesto la orientación de esta.

4.2 Sistema de detección de caracteres (Tesseract OCR)

A continuación, se realizan las pruebas utilizando las seis hojas anteriormente manejadas para medir la distancia, en este paso lo que se espera es conocer la cantidad de palabras que puede traducir el sistema desde la imagen capturada.

Figura 4.3 Imágenes tomadas por Pi Cámara V2.1



Fotos capturadas luego del reconocimiento del margen. Elaborado por: David Basantes, Ernesto Chalaco

Continuando con el proceso se puede ver en la Figura 4.4 los resultados de texto que se obtuvieron de cada una de las fotos en la Figura 4.3, donde se descubren las palabras reconocidas por el sistema.

Figura 4.4 Resultado de Tesseract OCR

[illegible]

Palabras reconocidas en cada una de las fotos. Elaborado por: David Basantes, Ernesto Chalaco.

Con estos resultados se realizan las operaciones correspondientes para obtener el porcentaje de eficiencia del sistema.

Tabla 4.2 Porcentajes de eficiencia del sistema.

Imagen	Tamaño de letra	Palabras en la Imagen	Palabras en el texto	Porcentaje de eficiencia [%]	Error
A	26	71	71	100	0.000
B	18	65	63	96.92	0.153
C	16	274	257	93.79	0.062
D	12	514	445	86.57	0.134
E	11	514	17	3.30	0.966
F	16-14-12	207	131	63.28	0.367
Total		597	467	78.22	0.217

Elaborado por: Ernesto Chalaco, David Basantes

En la Tabla 4.2 se puede observar el porcentaje individual y total de palabras acertadas con el programa, estas son las que TTV va a reconocer y pronunciar.

Se debe aclarar que el porcentaje de eficiencia del sistema depende de muchos factores como el ángulo de inclinación de la hoja o la distancia con respecto a las gafas, pero

dos factores importantes que se deben tomar en cuenta son: el nivel de luz, porque entre mayor sea la iluminación mejor será la definición que se obtiene en la imagen, y el segundo es el tamaño de letra, si es más grande, mayor cantidad de palabras traducidas se tendrá, esto concuerda con los resultados obtenidos donde a mayor tamaño de letra, se obtiene un error que se aproxima al 0%. Estas dos características son directamente proporcionales al producto final.

4.3 Validación del Prototipo

La validación del prototipo de gafas para discapacitados visuales se realizó con varias personas, quienes experimentaron su portabilidad y facilidad de funcionamiento. Se puede ver en la Figura 4.5 que el prototipo consta de la cámara adherida a las gafas en este caso, en el lado derecho de la persona, la cual se conecta a la Raspberry por medio de un cable flex, además de los audífonos conectados al jack de 3.5mm.

Figura 4.5 Prototipo de gafas de lectura para personas con discapacidad visual



Elaborado por: David Basantes, Ernesto Chalaco

El prototipo está destinado principalmente para el uso de guía y aprendizaje a través de la lectura de texto, de esta manera se puede enseñar a cualquier persona a usarlo, sin discriminar entre género y edades. Así en la Figura 4.6 se aprecia cómo está conectado el sistema y colocado en un usuario, con la cámara adherida a las gafas, los audífonos y por supuesto el sistema embebido de procesamiento.

Figura 4.6 Prototipo de gafas para personas con ceguera



Elaborado por: Ernesto Chalaco, David Basantes

Se puede ver en la Figura 4.6 la forma de uso del prototipo, de esta manera se realizó preguntas a 5 usuarios diferentes, en género y edad, para conocer 4 factores importantes en el diseño, estos son: el peso, la portabilidad, la aceptación y funcionalidad.

4.6.2 Porcentaje de validación del prototipo

Para obtener la validación del prototipo se realizó varias preguntas con una ponderación de 1 a 5, en donde 1 es el menor valor que se puede dar al prototipo y 5 es por consiguiente el mayor valor que puede obtener.

La primera pregunta, a la que se denomina Portabilidad, se realizó sobre el peso y la comodidad que tiene el prototipo, lo que se buscó en los usuarios es la sensación al tener en sus manos la tarjeta Raspberry Pi3 B+ conectada a su Power Pack, conjuntamente con la cámara adherida a las gafas, esto nos da una respuesta de molestia o comodidad sobre el sistema.

El segundo cuestionamiento, al que se denomina Utilidad, se buscó conocer el aporte que el prototipo tiene con el usuario, si este es, o no, una ayuda para las personas con discapacidad visual.

En tercer lugar, se tiene la Funcionalidad, con esta pregunta se buscó saber si los entrevistados tiene una dificultad alta para el manejo y desenvolvimiento del prototipo, se obtuvo opiniones sobre si la adaptación a su uso se da de una forma fácil o compleja.

Para la cuarta pregunta que la identificamos como Audio, los entrevistados tuvieron que identificar la voz sintetizada como entendible y clara, esto permitió saber si el audio de salida es lo suficientemente agradable para las personas.

La última pregunta se refiere a la Implementación, con ello se tuvo en cuenta si las personas que usaron el prototipo están de acuerdo en que, este puede ser una solución para el aprendizaje, ser replicada y comercializada para el público en general.

Todos los parámetros expuestos los podemos observar en la Tabla 4.3.

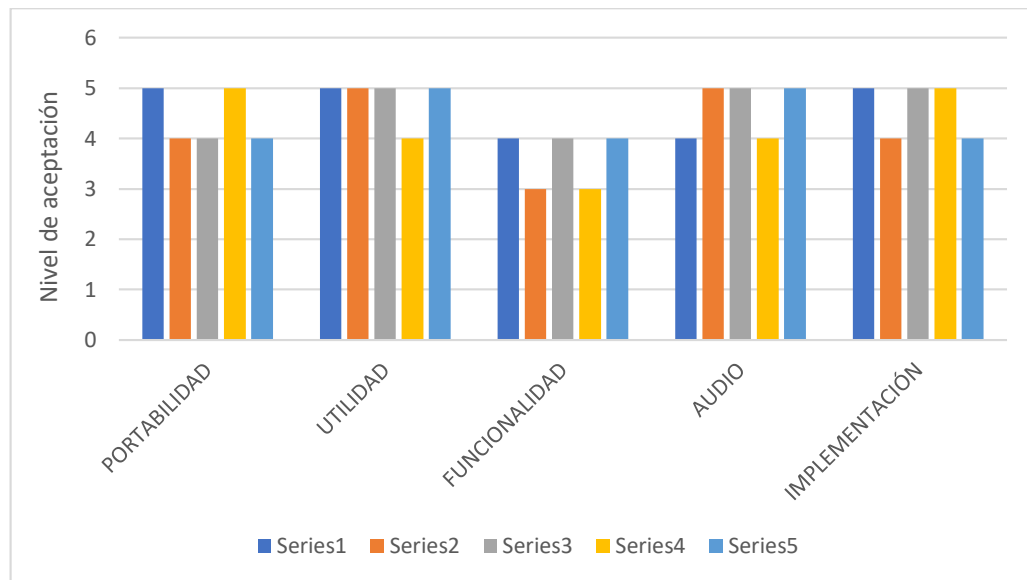
Tabla 4.3 Factores de validación del prototipo

PORTABILIDA D	UTILIDA D	FUNCIONALIDA D	AUDI O	IMPLEMENTACIÓ N	% VALIDACIÓ N
5	5	4	4	5	92
4	5	3	5	4	84
4	5	4	5	5	92
5	4	3	4	5	84
4	5	4	5	4	88

Elaborado por: Ernesto Chalaco, David Basantes

Además de realizar la encuesta, también es importante conocer la opinión más subjetiva de las personas, quienes mencionaron que la cámara adherida a las gafas, no representa una molestia, casi es imperceptible, también concordaron en la intromisión que representa el cableado que conecta la cámara, y mencionaron que debe ser más liviano y poco intrusivo para que no represente peso en los lentes. Para portabilidad los encuestados no tuvieron problema en tener en la mano el sistema embebido de procesamiento pero que es mejor si se tiene un estuche para ponerlo en el cinturón y poder llevarlo con mayor facilidad, finalmente algunos usuarios mencionaron que es una buena iniciativa que puede ser replicada y mejorada, accesible al público refiriéndose concretamente al factor económico.

Figura 4.7 Resultados de validación del prototipo



Validación de prototipo por factores, Elaborado por: Ernesto Chalaco, David Basantes

CONCLUSIONES

En la recopilación de información de papers y documentos sobre prototipos de gafas para personas con discapacidad visual existentes, se determinó que los componentes más adecuados para la construcción de un prototipo de bajo costo y sin que se afecte su funcionalidad son: Raspberry pi 3 B+, cámara PiCámara, batería de LiPo, audífonos con Jack de 3.5; debido a que en el estudio de características adecuadas son escogidos por su accesibilidad de compra y bajo costo, capacidad de procesamiento de la Raspberry pi 3 B+, voltaje, corriente y autonomía de la batería de LiPo y versatilidad a la hora de usarlos. Además, se constató que existen pocos modelos de gafas para discapacitados en el mercado, pero si muchos en desarrollo con diferentes funcionalidades y especificaciones.

El diseño del prototipo de las gafas para discapacitados se enfocó en la necesidad de exponer un método de visión artificial práctico, que ayude al desarrollo de un algoritmo capaz de reconocer un margen con las características en el proyecto mencionadas, que contenga el texto de interés de la persona no vidente, esto con el fin de darle mayor funcionalidad y aportar al desarrollo de nuevas tecnologías para discapacitados visuales con bajo costo. El prototipo al apoyarse en visión artificial logra disminuir un alto porcentaje de error, en el número de palabras traducidas por hoja, que en este caso puede ser debido a la mala ubicación de la hoja de texto, pasando de un 10% sin visión artificial a un 100 % de palabras traducidas con visión artificial.

La implementación del convertidor de texto a voz compatible con el sistema de visión artificial se la hace por medio del formato Text to Voice, manejada por la librería TTS (text to speech) de Python, esto debido a que tiene preinstalado voces sintetizadas que harán la función de lectores, para el prototipo se utilizó la voz en español de un hombre, ya que se consideró la utilización del prototipo en Ecuador y el tono masculino beneficia al entendimiento, además de tener menores inconvenientes de funcionalidad.

El uso de software gratuito utilizado en el desarrollo del sistema sumado a la facilidad de adquisición de los materiales utilizados en el proyecto dio lugar al abaratamiento

del prototipo (200 aproximadamente), tal es el caso de las gafas ORCAM para discapacitados con más aplicaciones, pero con un valor en el mercado superior a los 3000 dólares, haciéndolo de difícil adquisición en nuestro país. Para la base del sistema se usó OpenCv 4.0, que es la última versión que existe en la actualidad, con ello se implementó la librería Numpy propia de la distribución y que es específicas para usar con visión artificial, siendo esta la más importante de todas ya que permite realizar las transformaciones, arreglos de imagen necesarias para corregir cualquier defecto y mejorar las condiciones de la fotografía, con un resultado adecuado para su traducción a texto y posteriormente a voz.

Las pruebas de funcionamiento para la validación del prototipo se realizaron con diferentes tamaños y tipo de letra, en este último factor se usaron las más comunes: arial, times new roman y calibri, útiles para la traducción a texto plano. Se observa que a un tamaño de letra de 26 el error que se obtiene es prácticamente nulo, con el total de palabras traducidas y un error del 0%, viéndose afectado este rendimiento cuando llegamos a un tamaño de letra de 12, con un error de 13.4 %. Cuando se adquiere la fotografía los caracteres que están más cerca al punto central del enfoque de la lente tienen mayor posibilidad de ser traducidas y la detección de caracteres es más eficiente como consecuencia de la frontalidad que tiene el texto, por ello es adecuado tener una configuración de márgenes de al menos 5cm en cada lado de la hoja.

RECOMENDACIONES

Al trabajar con la PiCámara, se debe evitar hacer pruebas con movimiento bruscos que imposibiliten una adecuada captura del texto y por tanto una pésima traducción del texto. Cuando se manipula el Flex de la cámara esta debe ser lo suficientemente extensa como para permitir hacer uso y manejo del prototipo de manera fácil y sencilla.

El diseño y construcción del sistema tienen un alto porcentaje de aceptación por parte de las personas entrevistadas, siendo el factor más aceptado el peso al no ser un inconveniente de portabilidad y movilidad, pero el factor a tomar en cuenta es el flex de la PiCamara que se utilizó para la implementación del prototipo, porque, aunque tiene características de movilidad y portabilidad, generó un poco de incomodidad a los no videntes a la hora de utilizar las gafas.

Es necesaria la práctica al manejar el prototipo, para generar una manera intuitiva de utilizarlo, la practicidad se produce con el uso continuo, porque se establece la distancia, ubicación y tiempo de inmovilidad de la hoja de texto, logrando como resultado un bajo porcentaje de error en la traducción del texto.

REFERENCIAS

- Arráez, D. (06 de 10 de 2017). Unas gafas inteligentes para que los ciegos puedan ´ver´ con el oído. *Diario de Ibiza*, pág. 1.
- Caiza, G., & Garcia , M. (2017). Implementacion de sistemas distribuidos de bajo costo bajo norma IEC-61499. *Ingenius*, 3.
- Chen Wei-Ling, L. C.-H.-S.-C. (Enero de 2017). *Hypervolemia screening in predialysis healthcare for hemodialysis patients using fuzzy color reason analysis*. Obtenido de researchgate.net: https://www.researchgate.net/publication/312678134_Hypervolemia_screening_in_predialysis_healthcare_for_hemodialysis_patients_using_fuzzy_color_reason_analysis
- CONAFE, S. (2016). *Discapacidad visual - Guia didactica*. Mexico: CONAFE.
- Consejo Nacional para igualdad de discapacidades. (2019). personas con discapacidad registradas. *Consejo Nacional para igualdad de discapacidades*, 1.
- INEC. (Junio de 2019). *Instituto Nacional de Estadistica y Censos*. Obtenido de <http://www.ecuadorencifras.gob.ec/institucional/home/>
- Kultura. (2011). OCR: tecnología para el reconocimiento óptico de caracteres en una imagen. *Kultura*.
- Organizacion mundial de la salud. (2017). Ceguera y discapacidad visual. *Organizacion mundial de la salud*, 1.
- Ortí, C. B. (2012). LAS TECNOLOGÍAS DE LA INFORMACIÓN Y. *Universidad de valencia* , 7.
- OSL.ULL. (08 de 2015). *OpenCV: Librería de Visión por Computador*. Obtenido de <https://osl.ull.es/software-libre/opencv-libreria-vision-computador/>
- RO-BOTICA. (Junio de 2019). *RO-BOTICA*. Obtenido de <https://www.robotica.com/Producto/RASPBERRY-PI-3-MODELO-B/>
- Rodrigo, A. Á. (2014). *Sistema de sensorización haciendo uso de raspberry pi para su uso e implementación en un entorno inteligente*. España.
- Rubio, V. T. (2015). *Sistema de reconocimiento de matrículas y gestión de acceso en una plataforma*. Cartagena.
- Ruiz, C. A. (2017). Las Tecnologías de la Información y Comunicación para la inclusión. *IJERI*.
- Sanchez, Á. (2005). *Aplicaciones de la visión artificial y la biometría informatica*. Madrid: Dykinson.
- TECNOBILITY, D. P. (06 de 2019). *Tecnobility*. Obtenido de <https://www.tecnobility.com/es/noticia/un-anillo-que-lee-texto-para-personas-con-p%C3%A9rdida-de-visi%C3%B3n>
- Tendencias21. (Mayo de 2012). Tendencias tecnologicas.

VISION 2020. (2014). Cifras de Ceguera en Latinoamérica. *Vision 2020 Latinoamericana Boletín trimestral*, 1.